The Behavior Analyst 1988, 11, 71–75 No. 1 (Spring)

Laboratory Lore and Research Practices in the Experimental Analysis of Human Behavior: Use and Abuse of Subjects' Verbal Reports[1]

Michael Perone
West Virginia University

The experimental analysis of behavior traditionally has focused on direct observation of overt responses that are defined mechanically and recorded automatically. By comparison, research in other areas of psychology employs a broader range of observational devices, including several in which subjects act as observers of their own behavior. The relative lack of self-report data in the experimental analysis of human behavior can be traced to the influence of the nonhuman research that has dominated experimental analysis in general. In the decade ending in 1981, about 93 percent of empirical papers in the *Journal of the Experimental Analysis of Behavior* involved nonhuman subjects (Buskist & Miller, 1982)—a situation opposite that of other areas of psychology where human studies occupy ten times more journal space than nonhuman ones (Miller, 1985). Some authors have criticized the failure of behavior analysts to exploit the verbal capability of their subjects as "an excessive reliance on an animal model of human functioning" (Bentall, Lowe, & Beasty, 1985, p. 178). Whatever our judgement of it, the situation is clear: By comparison with general psychological research, behavior analytic research with humans has been quite restricted in the sorts of data it considers. Whereas other psychologists study a broad range of observations and make frequent use of subjects' self-reports, behavior analysts focus on directly observable instances of overt behavior.

VERBAL REPORTS AND BEHAVIOR ANALYSIS

Behavior analysts' interest in self-report data seems to be growing, however. I first became aware of this when Alan Baron and I submitted a paper to the *Journal of the Experimental Analysis of Behavior* (Perone & Baron, 1980). The work was concerned with conditioned reinforcement of human observing responses. We trained men on multiple schedules of monetary reinforcement and extinction; later the multiple-schedule stimuli were omitted unless the men performed observing responses that produced them. We found that the stimulus correlated with extinction functioned as a conditioned reinforcer, as did the stimulus correlated with reinforcement.

One of the reviewers commented as follows: "I was surprised that no verbal reports—summarizing what the subjects *said* they were doing and why—were presented. If available, they should be discussed." This struck me as an odd request. Why *ask* the subjects what they did, when the point of the study was to precisely measure what they did? If the subjects' reports had differed from our own observations, we would have had to reject the reports. On the other hand, if the reports had been consistent, little would have been added to our account of their behavior. The reviewer also suggested that we ask the subjects *why* they did as they did. Answering the question of why an instance of behavior occurs is the very business of experimental analysis. To the extent that subjects' behavior was under the control of the variables we manipulated, we believed that we already knew the answer.

Nevertheless, we could not resist asking a few questions at the end of the experiment. Remembering the classic controversy about whether awareness of contingencies is necessary for learning (Dulaney, 1968), we wondered whether the subjects could describe the contingencies to which they had been exposed.

[1] Reprints may be obtained from the author, Psychology Department, Box 6040, West Virginia University, Morgantown, WV 26506-6040.

The interviews were informal, open-ended, and thoroughly unsystematic, but nothing the subjects said was reinforcing enough to shape a more scientific approach. Frankly, most of the remarks were uninterpretable. The subjects usually failed to answer even simple questions about the functions of the colored lamps we used to signal the periods of reinforcement and extinction. This was surprising because the men worked at a highly repetitive task, in a relatively impoverished environment, inside a 3 × 3 foot cubicle, 4 hours a day, 5 days a week, for at least 3 weeks. Compared to the world outside the laboratory, very little could happen here. One might expect the subjects to be able to describe this environment in some detail, but they could not.

We also collected systematic data on the subjects' verbal reports, but not the kind of reports the journal reviewer had in mind. Our experiment was part of a larger project concerned with operant analyses of job performance and drug use among industrial workers. In addition to the subjects' behavior on operant schedules, we were interested in their less conspicuous reactions to the laboratory manipulations. We described the experiment to them as a job simulation. At regular intervals we gave standard tests to measure mood and job satisfaction and also to monitor drug use. To encourage candid replies, we gave the subjects envelopes in which to seal the questionnaires and promised that we would not open them until the end of the study.

Among other things, we found that men previously classified as drug users did report consumption of illicit drugs, whereas men classified as nonusers did not. As the 4-hour work day passed, subjects reported increasing levels of tension, depression, anger, fatigue and confusion, and decreasing levels of vigor. Perhaps most interesting, we found that reports of satisfaction with the simulated job were positively correlated with the subjects' satisfaction with their real jobs in industry. We interpreted our findings as demonstrating the external validity of laboratory studies, in that subjects reacted similarly to real and simulated jobs (Perone, DeWaard, & Baron, 1979).

Such data are, of course, limited. As verbal responses to questionnaires, their relationship to other aspects of behavior is uncertain. For example, one may ask about the degree to which job satisfaction scores correspond to a worker's absenteeism, productivity, or likelihood of filing a grievance. Furthermore, the responses were grouped for statistical analysis. Group averages, even statistically significant ones, do not necessarily portray the behavior of the individuals within the groups. Perhaps the greatest limitation was the correlational nature of the study. We did not take control of the subjects' reactions to the real and simulated jobs; at best we only measured them and noted certain regularities.

USE OF VERBAL REPORTS

Although generally neglected by behavior analysts, self-report data can be used in many ways. Four uses are especially prominent in the literature.

Observation of Behavior

First, self-report data may provide the only practical means of observing certain forms of behavior. In our correlational study, we asked subjects to report on their consumption of illicit drugs. Although such behavior is overt, it is not likely to occur in public and it would seem to be difficult to observe in the laboratory. It seems reasonable to ask subjects to report on instances of such behavior, but of course the way in which the reports are evoked is of paramount importance if reliability and validity are to be assured (cf. Baron & DeWaard, 1976). Collecting data on illegal and secretive activities such as drug use may be an extreme example, but the considerations described here apply as well to other naturally occurring behavior that takes place in private, as well as to subject characteristics that are difficult to ascertain directly, such as age, class standing, and health.

A second use of self-reports is to collect data on global characteristics of behavior

that essentially are defined in terms of verbal responses to a set of questions. Historically, this use of self-reports is associated with methodological behaviorism, and a clear example involves the measurement of attitudes. Operationally, an attitude toward some object is defined in terms of an average response to a set of statements about the object. In our correlational study, we administered standard tests of job satisfaction, that is, we measured attitudes toward the experimental task. A major limitation of this approach is uncertainty about the validity of the interpretations attached to the subjects' responses. When researchers label the scored responses to a questionnaire, they are placing an interpretation on the set of responses, and the resulting abstraction or construct—for example, "job satisfaction"—is sometimes treated as a functional unit. Whether such practices are justified is the question of "construct validity," and it is a difficult one. Behavior analysts perhaps can avoid this issue when the units of analysis in their studies are observable response classes. But to the extent that behavior analysts deal in abstractions—for example, "response strength" and "rule-governed behavior"—they too must grapple with the problem of construct validity (cf. Nunnally, 1978).

Verbal Reports as Primary Data

A third use for subjects' self-reports arises when the reports themselves constitute the object of study. This use stands in marked contrast to those described so far. In the previous cases, subjects observed and reported on their own behavior because the investigator was unwilling or unable to observe the behavior directly. By comparison, in the present case the subjects do not replace the experimenter as observers, for the verbal reports themselves are the responses of interest.

Two published studies meeting the above description have come from the University of Maryland Baltimore County (Catania, Matthews, & Shimoff, 1982; Matthews, Catania, & Shimoff, 1985). In the more recent study, college students earned money by pressing buttons on multiple random-interval random-ratio schedules and, after every exposure to the pair of schedules, by writing statements about either the contingencies or the appropriate way to press the buttons. Correct statements were shaped by reinforcing successive approximations with varying amounts of money. When subjects were trained to describe performances, their button-pressing generally was consistent with the descriptions—even if the descriptions contradicted the actual contingencies. By comparison, when subjects were trained to describe the contingencies, button pressing sometimes was and sometimes was not in accord with the descriptions. These data show that verbal behavior can interact with nonverbal behavior and that there are limits to the interaction: the pattern of button pressing paralleled verbal descriptions of performance but not descriptions of contingencies.

This line of research is original, and the procedures and analyses are, understandably, in need of refinement. Although central to the aims of the research, the shaping procedure and resulting verbal responses were not described as precisely as one would like. The experiments lasted just a few hours, so it is unclear how verbal and nonverbal behavior might interact after performances are allowed to stabilize. The inconsistent relation between contingency descriptions and button pressing suggests that identification of the controlling variables remains to be completed. Unfortunately, the procedures cannot isolate the contributions of several possible sources of control over the verbal descriptions, for example, the shaping procedure, the multiple schedule, the operant performances maintained by the schedule, or the interaction among these variables. Until such matters can be resolved, the relationship between the verbal descriptions and the schedule performances must be viewed essentially as correlational, not causal. Despite these

problems, Matthews et al.'s study represents an important step in the development of procedures in which verbal behavior can be studied as operant behavior, directly and objectively measured, and, most importantly, brought under explicit control of environmental contingencies of reinforcement.

Explanatory Use of Verbal Reports

The purpose of the fourth use of verbal reports seems to be to find a short-cut around the difficulties of a true experimental analysis. The general procedure consists of exposing subjects to schedules of reinforcement (often for just a few brief sessions), asking some questions, and explaining the schedule performance by correlating it with the answers to the questions. For example, Bentall et al. (1985) studied children of varying ages as they responded on fixed-interval schedules. Response patterns of the youngest subjects—infants less than a year old—resembled those of rats and pigeons, but patterns of the other subjects—children about 3 to 9 years old—did not. In accounting for this outcome, the authors observed that the infant subjects could not describe the schedules whereas the older ones could. They documented this by describing some of the childrens' utterances during the sessions and by quoting some of their answers to open-ended questions of the form "what makes the machine work?" Considering both the schedule performances and the verbal reports, Bentall et al. (1985) drew two conclusions. First, they said that children of 5 years or more produced verbal descriptions of the contingencies that functioned as rules governing their responding on the schedule. Second, and more generally, they argued that the results of this and earlier experiments suggest that "the development of verbal behavior greatly alters human operant performance and may account for many of the differences found between human and animal learning" (p. 165).

The key question is not whether this conclusion is correct or even reasonable, but rather whether it is justified on the basis of the data. As Shimoff (1984) has noted, "an experimental analysis of behavior generally seeks causes of behavior in the environment, not in other behavior. Verbal behavior may serve as an intermediate cause, as when it is part of an extended chain preceding some nonverbal response, but an experimental analysis will trace the chain to its environmental origins" (p. 1). When the laboratory behavior of human subjects differs from that of nonhumans, it is tempting to attribute the differences to the humans' complex verbal and cognitive abilities. More constructive, perhaps, would be to regard such failures in the same way as one would regard discrepancies among nonhuman studies—simply as instances in which the controlling environmental contingencies remain to be identified (cf. Baron & Perone, 1982).

But perhaps even this argument assumes more than it should. In questioning the usefulness of correlating verbal reports with other responses, we assume that a reasonable measure of verbal behavior has been presented. In Bentall et al.'s (1985) paper, the information about the subjects' verbal reports is highly selective and takes on the flavor of anecdotes and general impressions. Given the major role that the self-report data play in the theoretical interpretations, this is not enough. What is needed is a precise, quantitative, reliable, and valid instrument to collect and describe the reports. The eventual goal would be to develop measures of verbal behavior that are as refined as our measures of nonverbal behavior. At present, we may be too uncritical in accepting anecdotal verbal data.

There is an ironic twist to this last use of verbal reports. Interest in verbal reports has been defended on philosophical grounds by arguing that radical behaviorists recognize the importance of covert verbal events such as "self-instructions" and "self-rules" in human behavior, whereas only methodological behaviorists would restrict study to publicly observable events (e.g., see Bentall et al., 1985, pp. 178–179). Yet whatever their faults, methodological behaviorists have at least taken the business of collecting self-report data seriously (e.g., see Nun-

ally, 1978). They have developed a technology of test construction that ensures reliability, and they are acutely aware of the need for demonstrating the validity of their measures, even though behavior analysts might question their success at this. Concern with reliability and validity of verbal report measures is missing from the *Journal of the Experimental Analysis of Behavior,* although standards for nonverbal data are exacting. Perhaps we could learn something from the psychologists whose research has been based primarily on verbal reports.

CONCLUSION

Two uses of verbal reports appear to be of special significance to the experimental analysis of human behavior. One involves isolating the verbal report as an operant response to be carefully defined, directly observed, and brought under explicit experimental control. This approach, although in an early stage of development, holds promise. A different approach involves asking subjects to report on the nature of their nonverbal behavior and the variables influencing it. Little attention has been paid to the suitability of the methods for evoking these reports, and the reliability and validity of the data are uncertain. Even if these concerns can be laid to rest, the correlational nature of the data places narrow limits on the interpretations that can be supported.

Advances in our understanding of verbal influences on nonverbal behavior depend on research strategies that bring the verbal influences under the direct control of the experimenter. Procedures in which verbal stimuli are presented through experimenter-provided instructions have considerable merit in this regard, insofar as the characteristics of instructions can be manipulated in much the same way as other environmental variables (Baron & Galizio, 1983; Hayes & Brownstein, 1984). Considerably less adequate are procedures in which the subjects themselves are the source of stimuli, as in studies of so-called "self-instructions." Regardless of whether such behavior is inferred from interviews at the end of the experiment or from verbalizations during the course of the experiment, the critical stimuli are not under direct control. Thus, the procedures do not provide for a true experimental analysis and the results cannot bear an explanatory burden.

REFERENCES

Baron, A., & DeWaard, R. J. (1976, May). The interview as descriptive method in the operant analysis of drug use. Paper presented at the meeting of the Midwest Association of Behavior Analysis, Chicago, IL.

Baron, A., & Galizio, M. (1983). Instructional control in human operant behavior. *The Psychological Record, 33,* 495–520.

Baron, A., & Perone, M. (1982). The place of the human subject in the operant laboratory. *The Behavior Analyst, 5,* 143–158.

Bentall, R. P., Lowe, C. F., & Beasty, A. (1985). The role of verbal behavior in human learning: II. Developmental differences. *Journal of the Experimental Analysis of Behavior, 43,* 165–181.

Buskist, W., & Miller, H. L. (1982). The study of human operant behavior, 1958–1981: A topical bibliography. *The Psychological Record, 32,* 249–268.

Catania, A. C., Matthews, B. A., & Shimoff, E. (1982). Instructed versus shaped human verbal behavior: Interaction with nonverbal responding. *Journal of the Experimental Analysis of Behavior, 38,* 233–248.

Dulaney, D. E. (1968). Awareness, rules, and propositional control: A confrontation with S-R behavior theory. In T. R. Dixon & D. L. Horton (Eds.) *Verbal behavior and general behavior theory* (pp. 340–387). Englewood Cliffs, NJ: Prentice-Hall.

Hayes, S. C., & Brownstein, A. J. (1984). Verbal behavior: Is the human operant lab an ideal place to begin? *Experimental Analysis of Human Behavior Bulletin, 2,* 11–13.

Matthews, B. A., Catania, A. C., & Shimoff, E. (1985). Effects of uninstructed verbal behavior on nonverbal responding: Contingency descriptions. *Journal of the Experimental Analysis of Behavior, 43,* 155–164.

Miller, N. E. (1985). The value of behavioral research on animals. *American Psychologist, 40,* 423–440.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Perone, M., & Baron, A. (1980). Reinforcement of human observing behavior by a stimulus correlated with extinction or increased effort. *Journal of the Experimental Analysis of Behavior, 34,* 239–261.

Perone, M., DeWaard, R. J., & Baron, A. (1979). Satisfaction with real and simulated jobs in relation to personality variables and drug use. *Journal of Applied Psychology, 64,* 660–668.

Shimoff, E. (1984). Post-session questionnaires. *Experimental Analysis of Human Behavior Bulletin, 2,* 1.